Generalised gradients for virtual elements and applications to a posteriori error analysis

Théophile Chaumont-Frelet, Joscha Gedicke, Lorenzo Mascotto

Generalised gradients for virtual elements and applications to a posteriori error analysis

Théophile Chaumont-Frelet*, Joscha Gedicke†, Lorenzo Mascotto‡

Abstract

We rewrite the standard nodal virtual element method as a generalised gradient method. This re-formulation allows for computing a reliable and efficient error estimator by locally reconstructing broken fluxes and potentials. We prove the usual upper and lower bounds with constants independent of the stabilisation of the method and, under technical assumptions on the mesh, the degree of accuracy.

AMS subject classification: 65N12; 65N15.

Keywords: adaptivity; polygonal mesh; flux reconstruction; potential reconstruction; virtual element method; generalised gradient.

1 Introduction

The virtual element method (VEM) [1] is a generalisation of the finite element method (FEM) to polytopic meshes whose elements may have curved and/or hanging facets, internal cuts, and nonconvex shapes. Virtual element spaces typically consist of functions that solve local variational problems with polynomial data and are not required to be known in closed form, but rather via their evaluation through suitable degrees of freedom. Such degrees of freedom are selected so as to permit the computation of projection operators onto polynomial spaces, which are inserted in the bilinear form in order to preserve the polynomial consistency of the scheme. The well posedness of the method is guaranteed by correcting the projected bilinear form with a stabilisation, which has to be computable via the degrees of freedom.

One among the appealing features of the scheme is that adaptive mesh refinements admit the creation of hanging facets without the need of generating extra (useless) elements. The literature on the adaptive VEM is rather wide. We mention here only the early contributions [3,5,8]. Almost all the related references are concerned with residual-type error estimators, which present two major downsides:

- the constants in the reliability and efficiency bounds depend on the choice of the stabilisation;
- as in the finite element case [18], the efficiency constant depends on the degree of accuracy of the method [4].

The first issue becomes relevant when performing aggressive mesh refinement and/or coarsening: ad-hoc stabilisations should be designed so as to be robust with respect to small facets, highly distorted elements, and so on. Partial improvements are discussed in [2] where the upper bound is independent of the stabilisation for lowest order elements on triangular grids.

Using local flux equilibration strategies, p-robust error estimators were introduced for the FEM and discontinuous Galerkin methods; see, e.g., [6,16]. Inspired by the above works, a first attempt in deriving a p-robust error estimator for the VEM was performed in [11]. There, a p-robust error estimator was derived based on the computation of a virtual *global* flux. However, straightforwardly adapting the *localisation* procedure in [13, 15] to the VEM led to the loss of the efficiency, one of the main reasons being the lack of Galerkin orthogonality.

*Inria Univ. Lille and Laboratoire Paul Painlevé, 59655 Villeneuve-d'Ascq, France, theophile.chaumont@inria.fr

†Institut für Numerische Simulation, Universität Bonn, 53115 Bonn, gedicke@ins.uni-bonn.de

‡Dipartimento di Matematica e Applicazioni, Università di Milano Bicocca, 20125 Milan, Italy, lorenzo.mascotto@unimib.it; IMATI-CNR, 27100, Pavia, Italy; Fakultät für Mathematik, Universität Wien, 1090 Vienna, Austria

This motivated us to rewrite in this contribution the VEM as a generalised gradient method since the resulting formulation allows us to recover a suitable Galerkin orthogonality. More precisely, on each element K, we rewrite the usual virtual element discrete bilinear form, which typically looks like

$$a_h^K(\cdot,\cdot) = a^K(\Pi\cdot,\Pi\cdot) + S^K((I-\Pi)\cdot,(I-\Pi)\cdot),$$

being Π a computable polynomial projector and $S^K(\cdot,\cdot)$ a computable stabilisation, as a generalised gradient bilinear form. In particular, we shall write

$$a_h^K(\cdot,\cdot) = (\mathfrak{G}_h^K(\cdot),\nabla\cdot)_K,$$

where $\mathfrak{G}_h^K(\cdot)$ is a generalised gradient given by a piecewise Raviart-Thomas function over a subtriangulation of K, which contains information on the stabilisation. This generalised gradient is available for *any* computable stabilisation and is obtained by solving small, localised finite element problems. These problems can be independently solved in parallel and their solution is cheap.

The use of a generalised gradient is also instrumental in the design of a posteriori error estimators for interior penalty discontinuous Galerkin schemes on simplicial and tensor product meshes, which also suffer from a lack of Galerkin orthogonality caused by the stabilisation [15]. A similar construction was also used in [14] in the framework of approximation of solutions to nonlinear problems for skeletal methods.

This generalised gradient allows for the design of an error estimator that is reliable and efficient with bounds independent of the chosen stabilisation as well as the degree of accuracy of the method; see Theorem 5.3. Specifically, once the solution u_h and the generalised gradient $\mathfrak{G}_h(u_h)$ are known, the computation of the error estimator involves the solution to local primal and mixed finite element problems on vertex patches as defined in (5.5).

As such, the implementation of the error estimator can be made fully parallel and is no more expensive than computing a residual error estimator. As we detail in Remark 5.5 below, the constants involved in the upper and lower bounds generalise to polytopic meshes those obtained in the corresponding bounds for discontinuous Galerkin schemes in [15]: the resulting estimates are in a sense optimal, since they are as good as the state-of-the-art for another nonconforming numerical method on standard meshes.

Numerical examples show that the generalised gradient provides us with an approximation of the exact solution comparable to that obtained with a polynomial energy projection of the discrete solution, i.e., the standard error measure in virtual elements. These examples also illustrate that the estimator provides a fairly sharp upper bound on the error that does not deteriorates as p increases.

Outline of the paper. We introduce the setting and the necessary notation in Section 2. We introduce the VEM and an equivalent rewriting using a generalised gradient in Section 3. Some technical results are detailed in Section 4. In Section 5, we introduce a computable error estimator and state its key properties. The estimator consists of two terms, analysed in Sections 6 and 7, respectively. We present numerical examples in in Section 8. Finally, Appendix A presents technical results, which are related to the p-robustness of the efficiency estimate.

Main result. We introduce a generalised gradient for the VEM in Definition 3.3 and show that it can be used to equivalently rewrite the VEM in Theorem 3.4. The error estimator is introduced in (5.5), and its main properties are given in Theorem 5.3.

2 Setting

In Section 2.1, we introduce the functional spaces that we shall employ throughout. The model problem is given in Section 2.2, while polygonal meshes and their assumptions are detailed in Section 2.3. Several polynomial spaces are recalled in Section 2.4.

2.1 Functional spaces

Consider an open set D of $\mathbb{R}^2$. For q in $[1,\infty]$, we employ the standard notation $L^q(D)$ for Lebesgue spaces equipped with their usual norm $\|\cdot\|_{L^q(D)}$. When $q=2$, $L^2(D)$ is a Hilbert space, and its norm and inner-product of $L^2(D)$ are denoted by $\|\cdot\|_D$ and $(\cdot,\cdot)_D$. For vector-valued functions, we set $\boldsymbol{L}^2(D) := [L^2(D)]^2$, and use the same notation for its norm and inner-product. $H^1(D)$ and $\boldsymbol{H}(\mathrm{div}, D)$ stand for the Sobolev spaces of scalar-valued function v in $L^2(D)$ such that ∇v belongs to $\boldsymbol{L}^2(D)$ and vector-valued function $\boldsymbol{\tau}$ in $\boldsymbol{L}^2(D)$ with $\mathrm{div}\,\boldsymbol{\tau}$ in $L^2(D)$, respectively. $H^1_0(D)$ is the closure in $H^1(D)$ of the space of smooth functions with compact support in D. If $\mathcal{D}$ is finite a collection of disjoint open sets, then $H^1(\mathcal{D})$ collects functions v such that $v|_D$ is in $H^1(D)$ for all D in $\mathcal{D}$.

2.2 The continuous problem

We consider a fixed Lipschitz polygonal domain Ω and f in $L^2(\Omega)$. The model problem we are interested in consists in finding u in $H^1_0(\Omega)$ such that

$$(\nabla u, \nabla v)_\Omega = (f, v)_\Omega \qquad \forall v \in H^1_0(\Omega). \tag{2.1}$$

Standard arguments reveal that (2.1) is well posed. In what follows, u will always denote the solution to (2.1).

2.3 Regular polygonal meshes

We consider meshes $\mathcal{T}_h$ consisting of open, conforming polygonal elements over Ω. We denote their set of vertices and edges by $\mathcal{V}_h$ and $\mathcal{E}_h$. To each element K of $\mathcal{T}_h$, we associate its diameter h_K, centroid $\mathbf{x}_K$, set of vertices $\mathcal{V}^K$, and set of edges $\mathcal{E}^K$. We introduce the mesh size $h := \max_{K\in\mathcal{T}_h} h_K$ and the length h_e of each edge e in $\mathcal{E}_h$. We further split internal and boundary vertices into the sets $\mathcal{V}_h^I$ and $\mathcal{V}_h^B$, and internal and boundary edges into the sets $\mathcal{E}_h^I$ and $\mathcal{E}_h^B$. We denote the maximal number of vertices of the elements of $\mathcal{T}_h$ by N_{vert}.

Henceforth, we assume the following regularity properties to be valid: there exists γ in $(0,1]$ such that

- each element h_K is star-shaped with respect to a ball $B(K)$ of radius larger than or equal to γh_K;
- given an element K, any of its edges e has length h_e larger than γh_K.

The two above properties imply that each element K has a uniformly bounded number of vertices. Moreover, associated with each K, it is possible to construct a shape-regular subtriangulation $\widetilde{\mathcal{T}}_h^K$, by connecting the centre of the ball $B(K)$ with the vertices of K. Using the above regularity properties, we deduce the regularity of $\widetilde{\mathcal{T}}_h^K$. We define $\widetilde{\mathcal{T}}_h$ as the union of all $\widetilde{\mathcal{T}}_h^K$. Henceforth, we assume that such a subtriangulation $\widetilde{\mathcal{T}}_h$ is fixed. The broken gradient over either $\mathcal{T}_h$ or $\widetilde{\mathcal{T}}_h$ is denoted by ∇_h.

Only while deriving p-robust efficiency bound, we shall demand the further technical assumption on the mesh.

Assumption 2.1 (Single facet sharing)**.** *Any two distinct elements K and K' in $\mathcal{T}_h$ share at most one facet, i.e.,*

$$\partial K \cap \partial K' \in \mathcal{E}_h \qquad \forall K, K' \in \mathcal{T}_h,\ K \neq K',\ \partial K \cap \partial K' \neq \emptyset. \tag{2.2a}$$

For any boundary elements, we further have that

$$\partial K \cap \partial\Omega \in \mathcal{E}_h \qquad \forall K \in \mathcal{T}_h,\ \partial K \cap \partial\Omega \neq \emptyset. \tag{2.2b}$$

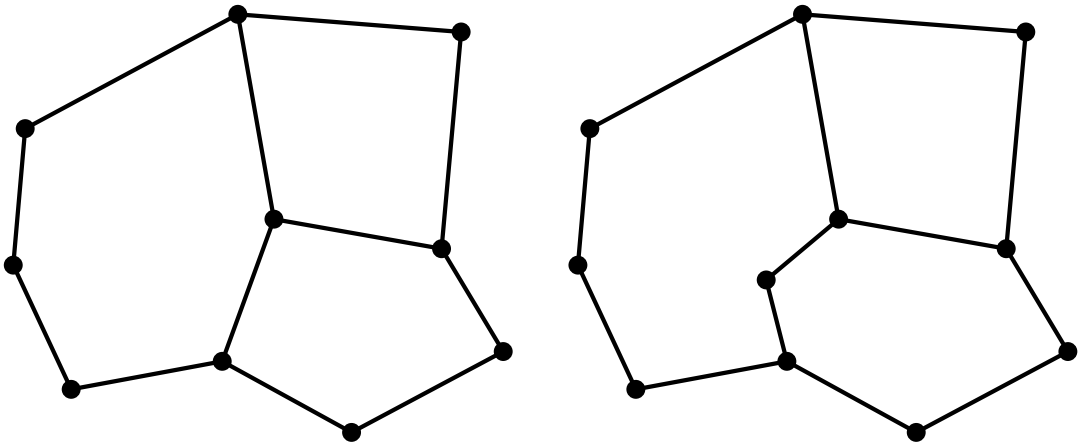

Figure 1: Examples of configurations allowed (*left panel*) and forbidden (*right panel*) under Assumption 2.1 for the derivation of p-robust estimates. The h-version allows for both configurations.

2.4 Standard polynomial spaces

If D in $\mathbb{R}^2$ is an open set and q is in $\mathbb{N}$, $\mathbb{P}_q(D)$ stands for the space of polynomials of degree at most q over D. $\mathbb{P}\!\!\mathbb{P}_q(D) := [\mathbb{P}_q(D)]^2$ contains vector-valued polynomials and $\mathbb{RT}_q(D) := \mathbb{P}\!\!\mathbb{P}_q(D) + \boldsymbol{x}\mathbb{P}_q(D)$ is the set of Raviart–Thomas polynomials. As before, if $\mathcal{D}$ is a collection of disjoint elements, then $\mathbb{P}_q(\mathcal{D})$, $\mathbb{P}\!\!\mathbb{P}_q(\mathcal{D})$, and $\mathbb{RT}_q(\mathcal{D})$ consist of piecewise polynomial functions over that collection.

3 The virtual element method and generalised gradients

We present two equivalent versions of the VEM. In Section 3.1, we introduce the virtual element spaces with their degrees of freedom, which allow us to compute stabilisations and projection operators in Section 3.2. The standard VEM is presented in Section 3.3, while an equivalent rewriting based on a generalised gradient is given in Section 3.4.

3.1 Virtual element spaces

We define [1] the virtual element space of order p in $\mathbb{N}$ on the element K as

$$V_h(K) := \left\{ v_h \in H^1(K) \mid \Delta v_h \in \mathbb{P}_{p-2}(K),\ \ v_{h|\partial K} \in \mathcal{C}^0(\partial K),\ v_{h|e} \in \mathbb{P}_p(e)\ \forall e \in \mathcal{E}^K \right\}.$$

We endow $V_h(K)$ with the following unisolvent degrees of freedom [1]: given v_h in $V_h(K)$,

- the point values of v_h at the vertices ν in $\mathcal{V}^K$;
- for each edge e of K, the point values of v_h at the $p-1$ internal Gauß-Lobatto nodes of e;
- given $\{m_{\boldsymbol{\alpha}}\}$ a basis of $\mathbb{P}_{p-2}(K)$ of elements that are invariant with respect to dilations and translations, the scaled moments

$$|K|^{-1} \int_K v_h\ m_{\boldsymbol{\alpha}} \qquad\qquad \forall |\boldsymbol{\alpha}| = 0, \dots, p-2.$$

The corresponding H^1 conforming global space reads

$$V_h := \{ v_h \in H^1(\Omega) \mid v_{h|K} \in V_h(K) \}.$$

A set of unisolvent degrees of freedom for V_h is given by the H^1 conforming coupling of the degrees of freedom of each $V_h(K)$.

3.2 Stabilisations and polynomial projectors

Consider the splitting

$$(\nabla u, \nabla v) = \sum_{K\in\mathcal{T}_h} (\nabla u, \nabla v)_K.$$

On each element K, we introduce a bilinear form $S^K(\cdot,\cdot)$ satisfying thee properties: (i) $S^K(1,1) > 0$; (ii) $S^K(\cdot,\cdot)$ is computable via the degrees of freedom of $V_h(K)$; (iii) there exist $0 < \alpha_* < \alpha^*$ such that

$$\alpha_* \left|v_h\right|^2_{1,K} \leq S^K(v_h, v_h) \leq \alpha^* \left|v_h\right|^2_{1,K} \qquad \forall v_h \in V_h(K), \quad (v_h, 1)_K = 0.$$

The degrees of freedom of $V_h(K)$ allow for the computation of projection operators onto polynomial spaces [1]. We define $\Pi_p^\nabla : V_h(K) \to \mathbb{P}_p(K)$ as

$$(\nabla(v_h - \Pi_p^\nabla v_h), \nabla q_p) = 0, \quad S^K(v_h - \Pi_p^\nabla v_h, 1) = 0 \qquad \forall v_h \in V_h(K),\ q_p \in \mathbb{P}_p(K). \tag{3.1}$$

The second condition above fixes constants due to the assumption $S^K(1,1) > 0$. Differently from the standard virtual element setting, we fix the average of $\Pi_p^\nabla v_h$ through the stabilisation $S^K(\cdot,\cdot)$. This will be instrumental while rewriting the method as a generalised gradient method in Section 3.4 below.

We define $\Pi^0_{p-2} : V_h(K) \to \mathbb{P}_{p-2}(K)$ as

$$(v_h - \Pi^0_{p-2} v_h, q_{p-2})_K = 0 \qquad \forall v_h \in V_h(K),\ q_{p-2} \in \mathbb{P}_{p-2}(K). \tag{3.2}$$

In what follows, with an abuse of notation, we introduce global projection operators Π_p^∇ and Π^0_{p-2} onto piecewise discontinuous polynomial spaces over $\mathcal{T}_h$ so that their restrictions to each element coincide with the operators in (3.1) and (3.2), respectively.

We introduce the local and global bilinear forms

$$\begin{aligned} a_h^K(\phi_h, v_h) &:= (\nabla(\Pi_p^\nabla \phi_h), \nabla(\Pi_p^\nabla v_h)) + S^K((I - \Pi_p^\nabla)\phi_h, (I - \Pi_p^\nabla)v_h), \\ a_h(\phi_h, v_h) &:= \sum_{K\in\mathcal{T}_h} a_h^K(\phi_h, v_h) \qquad \forall \phi_h, v_h \in V_h. \end{aligned}$$

In particular, the global discrete bilinear form $a_h(\cdot,\cdot)$ is coercive and continuous with respect to the energy norm with constants $\min(1,\alpha_*)$ and $\max(1,\alpha^*)$, respectively.

Remark 3.1. *Several explicit choices for the stabilisation $S^K(\cdot,\cdot)$ are available in the literature [17]. Amongst others, we mention the so-called "dofi-dofi" stabilisation given by the ℓ^2 inner product of the degrees of freedom and the "projected" stabilisation*

$$S^K(\phi_h, v_h) := h_K^{-2}(\Pi^0_{p-2}\phi_h, \Pi^0_{p-2} v_h)_K + h_K^{-1}(\phi_h, v_h)_{\partial K}.$$

At any rate, the forthcoming analysis is fairly general and works for any choice of $S^K(\cdot,\cdot)$ satisfying the three properties above.

3.3 The standard formulation of the VEM

We only consider the case where $p \geq 2$. The case $p = 1$ can be analogously handled with modifications as in [1]. Then, a virtual element method for the approximation of the solution to (2.1) reads

$$\begin{cases} \text{find } u_h \in V_h \text{ such that} \\ a_h(u_h, v_h) = (f, \Pi^0_{p-2} v)_\Omega \qquad \forall v_h \in V_h. \end{cases} \tag{3.3}$$

Method (3.3) is well posed due to the continuity and coercivity of $a_h(\cdot,\cdot)$, and the continuity of $(f, \Pi^0_{p-2}\cdot)_\Omega$. In what follows, u_h in V_h will always denote the solution to (3.3).

To simplify the presentation, we henceforth assume that f belongs to $\mathbb{P}_{p-2}(\mathcal{T}_h)$. This allows us to remove high-order data oscillation terms in the forthcoming bounds. The analysis of the general case poses no extra difficulties but would result in a more cumbersome notation.

3.4 The VEM as a generalised gradient method

We rewrite method (3.3) as a generalised gradient method. In other words, we describe how to design a suitable, fully-computable operator $\mathfrak{G}_h : V_h \to \boldsymbol{L}^2(\Omega)$ such that

$$a_h(\phi_h, v_h) = (\mathfrak{G}_h^K(\phi_h), \nabla v_h)_\Omega \qquad \forall \phi_h, v_h \in V_h. \tag{3.4}$$

We propose a construction that is fully-localised to each element, so that $\mathfrak{G}_h(\phi_h)|_K = \mathfrak{G}_h^K(\phi_h|_K)$ belongs to $\mathbb{RT}_p(\widetilde{\mathcal{T}}_h^K)$ and (3.4) holds true element-wise.

A key ingredient of the generalised gradient reconstruction is a "stabilisation lifting": for the explicit stabilisation

$$S^K(\cdot,\cdot) := h_K^{-2}(\Pi^0_{p-2}\cdot, \Pi^0_{p-2}\cdot)_K + h_K^{-1}(\cdot,\cdot)_{\partial K},$$

we can write

$$S^K((I-\Pi_p^\nabla)\phi_h, v_h) = (\mu_h^K(\phi_h), v_h)_{\partial E} - (r_h^K(\phi_h), v_h)_K, \tag{3.5}$$

where we have set

$$\mu_h^K(\phi_h) := h_K^{-1}\phi_h \in \mathbb{P}_p(\mathcal{E}^K), \qquad r_h^K(\phi_h) := -h_K^{-2}\Pi^0_{p-2}(I-\Pi_p^\nabla)\phi_h \in \mathbb{P}_{p-2}(K).$$

We show that it is *always* possible to rewrite the stabilisation as in (3.5), and that the data $\mu_h^K(\phi_h)$ and $r_h^K(\phi_h)$ can be found by solving a small linear system.

Lemma 3.2 (Stabilisation lifting)**.** *For all ϕ_h in $V_h(K)$, there exist unique $\mu_h^K(\phi_h)$ in $\mathbb{P}_p(\mathcal{E}^K)$ and $r_h^K(\phi_h)$ in $\mathbb{P}_{p-2}(K)$ such that*

$$S^K((I-\Pi_p^\nabla)v_h, \phi_h) = (\mu_h^K(v_h), \phi_h)_{\partial K} - (r_h^K(v_h), \phi_h)_K \qquad \forall v_h \in V_h(K). \tag{3.6}$$

Proof. The stabilisation is computable from the degrees of freedom; therefore, by Riesz' representation theorem, there exists a unique couple (r_h, μ_h) satisfying the above requirements.

In practice, consider the canonical basis $\{\varphi_j\}_{j=1}^{\dim(V_h(K))}$ of $V_h(K)$ dual to the degrees of freedom in Section 3.1 and $\mathbb{P}_p(\mathcal{E}^K)$ the space of piecewise continuous polynomials of maximum degree p over ∂K. This basis and a (polynomial) basis $\{q_\ell\}_{\ell=1}^{\dim(V_h(K))}$, $q_\ell = (q_\ell^{\partial K}, q_\ell^K)$, of the space $\mathbb{P}_p(\mathcal{E}^K)\times \mathbb{P}_{p-2}(K)$ have the same cardinality. In light of this, we solve the square system with coefficient matrix and right-hand side given by

$$\mathbf{A}_{i,j} := h_K^{-1}(q_j^{\partial K}, \varphi_i)_{\partial K} + h_K^{-2}(q_j^K, \varphi_i)_K, \quad \mathbf{b}_j := S^K((I-\Pi_p^\nabla)\phi_h, \varphi_j) \qquad \forall i,j = 1,\dots,\dim(V_h(K)).$$

We find r_h and μ_h by considering linear combinations with respect to the polynomial basis q_ℓ and coefficients given by the solutions to the above linear system. □

Next, we introduce the generalised gradient operator.

Definition 3.3 (Generalised gradient)**.** *For all K in $\mathcal{T}_h$ and ϕ_h in $V_h(K)$, we define*

$$\mathfrak{G}_h^K(\phi_h) := \nabla(\Pi_p^\nabla\phi_h - \mathcal{S}_h^K(\phi_h)) + \boldsymbol{\theta}_h^K(\phi_h) \in \mathbb{RT}_p(\widetilde{\mathcal{T}}_h^K), \tag{3.7}$$

where $\mathcal{S}_h^K(\phi_h)$ in $\mathbb{P}_p(K)$ satisfies

$$(\nabla\mathcal{S}_h^K(\phi_h), \nabla q_h)_K = S^K((I-\Pi_p^\nabla)\phi_h, q_h) \qquad \forall q_h \in \mathbb{P}_p(K)$$

and

$$\boldsymbol{\theta}_h^K(\phi_h) := \arg \min_{\substack{\boldsymbol{\tau}_h \in \mathbb{RT}_{p+1}(\widetilde{\mathcal{T}}_h^K)\cap \boldsymbol{H}(\mathrm{div},K) \\ \boldsymbol{\tau}_h\cdot\boldsymbol{n} = \mu_h^K(\phi_h) \\ \mathrm{div}\,\boldsymbol{\tau}_h = r_h^K(\phi_h)}} \|\boldsymbol{\tau}_h\|_K. \tag{3.8}$$

If ϕ_h belongs to V_h, then we define $\mathfrak{G}_h(\phi_h)$ in $\mathbb{RT}_p(\widetilde{\mathcal{T}}_h)$ by setting $\mathfrak{G}_h(\phi_h)|_K := \mathfrak{G}_h^K(\phi_h|_K)$.

The compatibility conditions for the definition of $\boldsymbol{\theta}_h^K$ in (3.8) are guaranteed by the way we fix the constant part of Π_p^∇ in (3.1) and identity (3.6).

The operator $\mathfrak{G}_h$ introduced in (3.7) allows for an equivalent rewriting of the VEM.

Theorem 3.4 (Rewriting of the VEM)**.** *Let K in $\mathcal{T}_h$. For all ϕ_h and v_h in $V_h(K)$, we have*

$$a_h^K(\phi_h, v_h) = (\mathfrak{G}_h^K(\phi_h), \nabla v_h)_K.$$

Similarly, if ϕ_h, v_h belong to V_h, then

$$a_h(\phi_h, v_h) = (\mathfrak{G}_h(\phi_h), \nabla v_h)_\Omega \qquad \forall v_h \in V_h.$$

Proof. Let ϕ_h, v_h be in $V_h(K)$. We split the local discrete bilinear form $a_h^K(\cdot,\cdot)$ into three terms:

$$a_h^K(\phi_h, v_h) = (\nabla(\Pi_p^\nabla \phi_h), \nabla(\Pi_p^\nabla v_h))_K + S^K((I - \Pi_p^\nabla)\phi_h, v_h) - S^K((I - \Pi_p^\nabla)\phi_h, \Pi_p^\nabla v_h).$$

Since Π_p^∇ is an orthogonal projection, we have

$$(\nabla(\Pi_p^\nabla \phi_h), \nabla(\Pi_p^\nabla v_h))_K = (\nabla(\Pi_p^\nabla \phi_h), \nabla v_h)_K$$

and, by the definition of $\mathcal{S}_h := \mathcal{S}_h^K(\phi_h)$,

$$S^K((I - \Pi_p^\nabla)\phi_h, \Pi_p^\nabla v_h) = (\nabla \mathcal{S}_h, \nabla(\Pi_p^\nabla v_h))_K = (\nabla \mathcal{S}_h, \nabla v_h)_K.$$

This leads us to

$$a_h^K(\phi_h, v_h) = (\nabla(\Pi_p^\nabla \phi_h - \mathcal{S}_h), \nabla v_h)_K + S^K((I - \Pi_p^\nabla)\phi_h, v_h).$$

For the remaining term, we introduce $\boldsymbol{\theta}_h := \boldsymbol{\theta}_h^K(\phi_h)$ and invoke Lemma 3.2, giving

$$\begin{aligned} S^K((I - \Pi_p^\nabla)\phi_h, v_h) &= -(r_h^K(\phi_h), v_h)_K + (\mu_h^K(\phi_h), v_h)_{\partial K} \\ &= -(\operatorname{div} \boldsymbol{\theta}_h, v_h)_K + (\boldsymbol{\theta}_h \cdot \boldsymbol{n}, v_h)_{\partial K} = (\boldsymbol{\theta}_h, \nabla v_h)_K, \end{aligned}$$

thereby concluding the proof. □

4 Technical results

The aim of this section is to discuss some technical results, which are useful in the a posteriori error analysis in Section 5, 6, and 7 below. In Section 4.1, we introduce a virtual partition of unity and analyse its properties. In Section 4.2, we investigate stability properties of certain minimisation problems.

4.1 Virtual element partition of unity

We introduce a virtual element partition of unity over the mesh $\mathcal{T}_h$, which we shall employ below in the localisation of the computation of the error estimator. For each vertex ν of $\mathcal{V}_h$, we define φ^ν as the only function in V_h that is harmonic in each element, piecewise linear on the skeleton of $\mathcal{T}_h$, equal to 1 at the vertex ν, and equal to 0 at all other vertices. This set of functions is a partition of unity; see, e.g., [19, Section 2.1].

For ν in $\mathcal{V}_h$, we denote the set of K in $\mathcal{T}_h$ sharing the vertex ν by $\mathcal{T}_h^\nu$. Then, $\omega^\nu := \operatorname{supp} \varphi^\nu$ corresponds to the set covered by the elements in $\mathcal{T}_h^\nu$. We denote the set of edges sharing the vertex ν by $\mathcal{E}_{h,\nu} \subset \mathcal{E}_h$ and, if ν is a boundary vertex, its subset of (two) edges lying on the boundary $\partial\Omega$ by $\mathcal{E}_{h,\nu}^{\mathrm{B}}$.

The space $\widetilde{H}^1(\omega^\nu) \subset H^1(\omega^\nu)$ defined for ν in $\mathcal{V}_h^I$ as

$$\widetilde{H}^1(\omega^\nu) := \left\{ v \in H^1(\omega^\nu) \mid (v,1)_{\omega^\nu} = 0 \right\} \tag{4.1a}$$

and for ν in $\mathcal{V}_h^B$ as

$$\widetilde{H}^1(\omega^\nu) := \left\{ v \in H^1(\omega^\nu) \mid v_{|e} = 0 \quad \forall e \in \mathcal{E}_{h,\nu}^{\mathrm{B}} \right\} \tag{4.1b}$$

will be useful later on, as well as the broken counterpart defined by $\widetilde{H}^1(\mathcal{T}_h^\nu) := H^1(\mathcal{T}_h^\nu)$ and $\widetilde{H}^1(\mathcal{T}_h^\nu) := \{v \in H^1(\mathcal{T}_h^\nu) \mid (v,1)_{\omega^\nu} = 0\}$ for interior and boundary vertices, respectively.

We introduce the average operator over any edge e in $\mathcal{E}_h$:

$$\Pi_0^{0,e} v = h_e^{-1}(v,1)_e \qquad \forall v \in L^1(e).$$

Remark 4.1. *In the case of simplicial meshes with finite element and discontinuous Galerkin methods [15], the standard partition of unity given by hat functions φ^ν satisfies some standard properties. More precisely, the scaling properties*

$$\|\varphi^\nu\|_{L^\infty(\Omega)} = 1, \qquad \|\nabla\varphi^\nu\|_{L^\infty(K)} \le C(\gamma)h_K^{-1} \quad \forall K \in \mathcal{T}_h, \tag{4.2}$$

combined with the product rule and a broken Poincaré inequality as in, e.g., [7], yield the inequality

$$\|\nabla_h(\varphi^\nu\theta)\|_{\omega^\nu}^2 \le C_{\mathrm{cont},\nu}^2(\|\nabla_h\theta\|_{\omega^\nu}^2 + \sum_{e\in\mathcal{E}_{h,\nu}} h_e^{-1}\|\Pi_0^{0,e}[\![\theta]\!]\|_e^2) \qquad \forall\theta \in \widetilde{H}^1(\mathcal{T}_h^\nu), \tag{4.3}$$

where the constant $C_{\mathrm{cont},\nu}$ only depends on γ.

We establish corresponding properties as those in Remark 5.5 for the case of polytopic meshes. We have the following result generalising (4.2).

Lemma 4.2. *For every ν in $\mathcal{V}_h$, the partition of unity function φ^ν satisfies*

$$\|\varphi^\nu\|_{L^\infty(\Omega)} = 1, \qquad \|\nabla\varphi^\nu\|_{L^4(K)} \le C(\gamma)h_K^{-\frac{1}{2}} \qquad \forall K \in \mathcal{T}_h. \tag{4.4}$$

Proof. The identity in (4.4) follows from the maximum principle for harmonic functions and the fact that φ^ν is piecewise linear over the skeleton of the mesh, equal to 1 at the vertex ν, and 0 at all other vertices.

As for the bound in (4.4), we first recall the scaled continuous Sobolev embedding $H^{\frac{1}{2}}(K) \hookrightarrow L^4(K)$ given by

$$\|\nabla\varphi^\nu\|_{L^4(K)} \le C(\gamma)\left(h_K^{-\frac{1}{2}}\|\nabla\varphi^\nu\|_K + |\nabla\varphi^\nu|_{\frac{1}{2},K}\right).$$

Since φ^ν is the solution to a Laplace problem with piecewise continuous linear polynomial Dirichlet boundary conditions, we use standard a priori estimates for elliptic problems, a (lowest order) polynomial inverse inequality on ∂K, and the fact that the mesh has no "small edges", and deduce

$$\|\nabla\varphi^\nu\|_{L^4(K)} \le C(\gamma)h_K^{-1}\|\varphi^\nu\|_{\partial K} \le h_K^{-1}\|\varphi^\nu\|_{L^\infty(\partial K)}|\partial K|^{\frac{1}{2}} = h_K^{-1}|\partial K|^{\frac{1}{2}} \le C(\gamma)h_K^{-\frac{1}{2}}.$$

□

We also have the following result generalising (4.3), which follows from the chain rule, the piecewise Sobolev embeddings $H^1(T) \hookrightarrow L^4(T)$ over the subtriangulation, Lemma 4.2, and a broken Poincaré inequality [7] on the subtriangulation.

Corollary 4.3. *For every ν in $\mathcal{V}_h$, there exists a positive constant $C_{\mathrm{cont},\nu}$ only depending on γ such that*

$$\|\nabla_h(\varphi^\nu\theta)\|_{\omega^\nu}^2 \le C_{\mathrm{cont},\nu}^2(\|\nabla_h\theta\|_{\omega^\nu}^2 + \sum_{e\in\mathcal{E}_{h,\nu}} h_e^{-1}\|\Pi_0^{0,e}[\![\theta]\!]\|_e^2) \qquad \forall\theta \in \widetilde{H}^1(\widetilde{\mathcal{T}}_h^\nu). \tag{4.5}$$

4.2 Discrete stable minimisation

In the proof of the efficiency of the error estimator, we shall need two technical bounds: for each vertex ν in $\mathcal{V}_h$, there exists a constant $C_{\mathrm{stab},\nu}$ such that

$$\min_{v_h\in\mathbb{P}_{p+2}(\widetilde{\mathcal{T}}_h^\nu)\cap\widetilde{H}^1(\omega^\nu)} \|\boldsymbol{G}_h - \nabla v_h\|_{\omega^\nu} \le C_{\mathrm{stab},\nu} \min_{v\in\widetilde{H}^1(\omega^\nu)} \|\boldsymbol{G}_h - \nabla v\|_{\omega^\nu} \tag{4.6a}$$

and

$$\min_{\substack{\boldsymbol{\tau}_h\in\mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu)\cap H(\mathrm{div},\omega^\nu)\\ \mathrm{div}\,\boldsymbol{\tau}_h = r_h}} \|\boldsymbol{G}_h + \boldsymbol{\tau}_h\|_{\omega^\nu} \le C_{\mathrm{stab},\nu} \min_{\substack{\boldsymbol{\tau}\in H(\mathrm{div},\omega^\nu)\\ \mathrm{div}\,\boldsymbol{\tau} = r_h}} \|\boldsymbol{G}_h + \boldsymbol{\tau}\|_{\omega^\nu} \tag{4.6b}$$

for all $\boldsymbol{G}_h$ in $\mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu)$ and r_h in $\mathbb{P}_p(\widetilde{\mathcal{T}}_h^\nu)$.

Under Assumption 2.1, in Appendix A.2, we show that $C_{\mathrm{stab},\nu}$ depends on the shape-regularity parameter γ but not on the degree of accuracy p. In the general case, in Appendix A.1, we show that (4.6) holds true with a constant depending on γ and possibly on p. We conjecture that the dependence on p for general meshes is artificial and could be removed at the price of a more technical proof.

5 A posteriori error estimator

The goal of this section is to present the general framework on which the estimator is based and propose an intuitive motivation for its definition. We describe how the estimator can be efficiently computed in practice, and state its key reliability and efficiency properties in Theorem 5.3. The actual proof of Theorem 5.3 is postponed to Sections 6 and 7 below.

The remainder of the section is organised as follows. We introduce the concept of generalised gradient in Section 5.1 and use it to derive a generalised Prager-Synge identity in Section 5.2. We introduce the error estimator and motivate its structure in Section 5.3, discuss its practical computability in Section 5.4, and state the main result of the paper, namely Theorem 5.3, along with some comments in Section 5.5.

5.1 Generalised gradient

The result of this section holds true for any generalised gradient satisfying Assumption 5.1 below. In particular, the generalised gradient $\mathfrak{G}_h(u_h)$ proposed in Section 3.4 satisfies this assumption.

Assumption 5.1 (Generalised gradient)**.** *$\boldsymbol{\mathcal{G}}_h$ in $\mathbb{RT}_p(\widetilde{\mathcal{T}}_h)$ is a generalised gradient if*

$$(\boldsymbol{\mathcal{G}}_h, \nabla\varphi^\nu)_\Omega = (f, \varphi^\nu)_\Omega \qquad \forall \nu \in \mathcal{V}_h^I.$$

5.2 A generalised Prager-Synge identity

The construction of the error estimator hinges upon a generalised Prager–Synge identity, as introduced in [15] for discontinuous Galerkin methods. For completeness, we adapt the proof of [15, Theorem 3.3] to the current setting.

Theorem 5.2 (Prager–Synge)**.** *For all $\boldsymbol{\mathcal{G}}_h \in \boldsymbol{L}^2(\Omega)$, we have*

$$\|\nabla u - \boldsymbol{\mathcal{G}}_h\|_\Omega^2 = \min_{s\in H_0^1(\Omega)} \|\boldsymbol{\mathcal{G}}_h - \nabla s\|_\Omega^2 + \sup_{\substack{v\in H_0^1(\Omega)\\ \|\nabla v\|_\Omega=1}} \left\{(f,v)_\Omega - (\boldsymbol{\mathcal{G}}_h, \nabla v)_\Omega\right\}^2. \tag{5.1}$$

Proof. Let $s^\star$ be the orthogonal projection of $\boldsymbol{\mathcal{G}}_h$ onto $H_0^1(\Omega)$, i.e.,

$$(\nabla s^\star, \nabla v) = (\boldsymbol{\mathcal{G}}_h, \nabla v) \qquad \forall v \in H_0^1(\Omega). \tag{5.2}$$

Pythagoras' theorem gives

$$\|\nabla u - \boldsymbol{\mathcal{G}}_h\|_\Omega^2 = \|\boldsymbol{\mathcal{G}}_h - \nabla s^\star\|_\Omega^2 + \|\nabla(u - s^\star)\|_\Omega^2 = \min_{s\in H_0^1(\Omega)} \|\boldsymbol{\mathcal{G}}_h - \nabla s\|_\Omega^2 + \|\nabla(u - s^\star)\|_\Omega^2.$$

Using that u is the solution to (2.1) and definition (5.2), the conclusion is a consequence of the following identities:

$$\begin{aligned}\|\nabla(u - s^\star)\|_\Omega &= \sup_{\substack{v\in H_0^1(\Omega)\\ \|\nabla v\|_\Omega=1}} (\nabla(u - s^\star), \nabla v)_\Omega \\ &= \sup_{\substack{v\in H_0^1(\Omega)\\ \|\nabla v\|_\Omega=1}} \{(f,v)_\Omega - (\nabla s^\star, \nabla v)_\Omega\} = \sup_{\substack{v\in H_0^1(\Omega)\\ \|\nabla v\|_\Omega=1}} \{(f,v)_\Omega - (\boldsymbol{\mathcal{G}}_h, \nabla v)_\Omega\}.\end{aligned}$$

□

5.3 Motivation and definition of the error estimator

We can employ duality in optimisation to control the last term on the right-hand side of (5.1) as

$$\sup_{\substack{v\in H_0^1(\Omega)\\ \|\nabla v\|_\Omega=1}} \{(f,v)_\Omega - (\boldsymbol{\mathcal{G}}_h, \nabla v)_\Omega\} \le \min_{\substack{\boldsymbol{\sigma}\in H(\mathrm{div},\Omega)\\ \mathrm{div}\,\boldsymbol{\sigma}=f}} \|\boldsymbol{\mathcal{G}}_h + \boldsymbol{\sigma}\|_\Omega.$$

Therefore, identity (5.1) states that the error is bounded by the sum of two contributions measuring the lack of $H^1_0(\Omega)$ and $\boldsymbol{H}(\mathrm{div},\Omega)$ conformity of $\boldsymbol{\mathcal{G}}_h$, i.e.,

$$\|\nabla u - \boldsymbol{\mathcal{G}}_h\|_\Omega^2 \leq \min_{s\in H^1_0(\Omega)} \|\boldsymbol{\mathcal{G}}_h - \nabla s\|_\Omega^2 + \min_{\substack{\boldsymbol{\sigma}\in\boldsymbol{H}(\mathrm{div},\Omega)\\ \mathrm{div}\,\boldsymbol{\sigma}=f}} \|\boldsymbol{\mathcal{G}}_h + \boldsymbol{\sigma}\|_\Omega^2. \tag{5.3}$$

In the construction of the error estimator, we want to measure these lacks of conformity using local computations on vertex patches. In other words, we expect that

$$\|\nabla u - \boldsymbol{\mathcal{G}}_h\|_\Omega^2 \simeq \sum_{\nu\in\mathcal{V}_h} \left\{ \min_{s^\nu\in\widetilde{H}^1(\omega^\nu)} \|\boldsymbol{\mathcal{G}}_h - \nabla s^\nu\|_{\omega^\nu}^2 + \min_{\substack{\boldsymbol{\sigma}^\nu\in H(\mathrm{div},\omega^\nu)\\ \mathrm{div}\,\boldsymbol{\sigma}^\nu=f}} \|\boldsymbol{\mathcal{G}}_h + \boldsymbol{\sigma}^\nu\|_{\omega^\nu}^2 \right\} \tag{5.4}$$

in some suitable sense, where $\widetilde{H}^1(\omega^\nu)$ is defined in (4.1a). This idea of "broken equilibration" was already partially explored in [9] but only for the second term.

The minimisation problems on the right-hand side are local but still not fully computable. Due to the discrete stable minimisations in Section 4.2, the proposed error estimator is based on replacing the minimisation problems with their natural finite element counterparts, namely,

$$\eta_{\nu,\mathrm{PT}} := \min_{v_h\in\widetilde{H}^1(\omega^\nu)\cap\mathbb{P}_p(\widetilde{\mathcal{T}}_h^\nu)} \|\boldsymbol{\mathcal{G}}_h - \nabla v_h\|_{\omega^\nu} \tag{5.5a}$$

and

$$\eta_{\nu,\mathrm{FL}} := \min_{\substack{\boldsymbol{\tau}_h\in H(\mathrm{div},\omega^\nu)\cap\mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu)\\ \mathrm{div}\,\boldsymbol{\tau}_h=f}} \|\boldsymbol{\mathcal{G}}_h + \boldsymbol{\tau}_h\|_{\omega^\nu}. \tag{5.5b}$$

The minimisers, which we shall henceforth denote by s_h^ν and $\boldsymbol{\sigma}_h^\nu$, are computable as solution to local finite element problems, so that the proposed estimator is fully computable.

5.4 Practical computation of the estimator on vertex patches

To compute $\eta_{\nu,\mathrm{PT}}$ in (5.5a), we solve

$$\begin{cases} \text{find } s_h^\nu \in \mathbb{P}_{p+2}(\widetilde{\mathcal{T}}_h^\nu)\cap\widetilde{H}^1(\omega^\nu) \text{ such that} \\ (\nabla s_h^\nu, \nabla z_h)_{\omega^\nu} = (\boldsymbol{\mathcal{G}}_h, \nabla z_h)_{\omega^\nu} \qquad \forall z_h \in \mathbb{P}_{p+2}(\widetilde{\mathcal{T}}_h^\nu)\cap\widetilde{H}^1(\omega^\nu), \end{cases} \tag{5.6}$$

where the mean-value may be fixed arbitrarily for internal vertices, since only ∇s_h^ν appears in the definition of the estimator.

To compute $\eta_{\nu,\mathrm{FL}}$ in (5.5b), we have to compute the discrete minimiser in (5.5b), i.e., the discrete flux solving

$$\begin{cases} \text{find } \boldsymbol{\sigma}_h^\nu \in \mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu)\cap\boldsymbol{H}(\mathrm{div},\omega^\nu) \text{ and } r_h^\nu\in\mathbb{P}_p(\widetilde{\mathcal{T}}_h^\nu) \text{ such that} & \\ (\boldsymbol{\sigma}_h^\nu, \boldsymbol{\tau}_h^\nu)_{\omega^\nu} + (\mathrm{div}\,\boldsymbol{\tau}_h^\nu, r_h^\nu) = (-\boldsymbol{\mathcal{G}}_h, \boldsymbol{\tau}_h^\nu)_{\omega^\nu} & \forall \boldsymbol{\tau}_h^\nu \in \mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu)\cap\boldsymbol{H}(\mathrm{div},\omega^\nu) \\ (\mathrm{div}\,\boldsymbol{\sigma}_h^\nu, q_h^\nu)_{\omega^\nu} = (f, q_h^\nu)_{\omega^\nu} & \forall q_h^\nu \in \mathbb{P}_p(\widetilde{\mathcal{T}}_h^\nu). \end{cases} \tag{5.7}$$

The above finite element problems are local on each vertex patch and can be solved in parallel.

5.5 Reliability and efficiency

For all ν in $\mathcal{V}_h$, we define the vertex estimator as

$$\eta_\nu^2 := \eta_{\nu,\mathrm{FL}}^2 + \eta_{\nu,\mathrm{PT}}^2 + \|\boldsymbol{\mathcal{G}}_h - \nabla_h(\Pi_p^\nabla u_h)\|_{\omega^\nu}^2 + \sum_{e\in\mathcal{E}_{h,\nu}} h_e^{-1}\|\Pi^0[\![\Pi_p^\nabla u_h]\!]\|_e^2$$

and the global error estimator as

$$\eta^2 := \sum_{\nu\in\mathcal{V}_h} \eta_\nu^2. \tag{5.8}$$

For a generic union D of general polygonal elements, we shall be interested in the error measure

$$\mathcal{E}^2(D) := \|\nabla u - \boldsymbol{\mathcal{G}}_h\|_D^2 + \|\boldsymbol{\mathcal{G}}_h - \nabla_h(\Pi_p^\nabla u_h)\|_D^2 + \sum_{\substack{e\in\mathcal{E}_h\\ e\subset D}} h_e^{-1}\|\Pi^0[\![\Pi_p^\nabla u_h]\!]\|_e^2. \tag{5.9}$$

In Section 8.1 below, we shall check the performance of (5.9) and compare it with a more standard choice: the two measures behave essentially the same on different meshes and for different degrees of accuracy.

Theorem 5.3. *Let γ be the shape-regularity constant introduced in Section 2.3; N_{vert} the maximal number of vertices per element in the mesh; $C_{\mathrm{cont},\nu}$ as in (4.5); $C_{\mathrm{stab},\nu}$ as in (4.6). The following reliability and local efficiency estimates hold true:*

$$\begin{aligned}
&\mathcal{E}^2(\Omega) \le 4N_{\mathrm{vert}}\sum_\nu C_{\mathrm{cont},\nu}^2\eta_\nu^2 \le 4N_{\mathrm{vert}}\max_{\nu\in\mathcal{V}_h} C_{\mathrm{cont},\nu}^2\sum_\nu \eta_\nu^2 =: c_{\mathrm{rel}}\eta^2,\\
&\eta_\nu \le 2C_{\mathrm{stab},\nu}^2\mathcal{E}^2(\omega^\nu) =: c_{\mathrm{eff}}\mathcal{E}(\omega^\nu).
\end{aligned} \tag{5.10}$$

The constants c_{rel} and c_{eff} are positive, independent of the choice of the stabilisation and the mesh size, but dependent on γ. The constant c_{eff} possibly depends also on the degree of accuracy p through $C_{\mathrm{stab},\nu}$. If two distinct mesh elements only share one facet as stated in Assumption 2.1, then c_{eff} is independent of p.

We believe that c_{eff} in (5.10) is independent of p also for meshes, which do not satisfy Assumption 2.1.

The computation of the error estimator involves the following quantities:

- the generalised gradient $\boldsymbol{\mathcal{G}}_h$ in (3.7), which can be seen as an elementwise post-processing of the VE solution;
- the solution of primal finite element problems on vertex patches as in (5.6) for the computation of $\eta_{\nu,\mathrm{PT}}$;
- the solution of mixed finite element problems on vertex patches as in (5.7) for the computation of $\eta_{\nu,\mathrm{FL}}$.

The computation of the error estimator is local and parallelizable, whence its cost is comparable to that of a residual error estimator. Furthermore, on meshes satisfying Assumption 2.1, it is p-robust and the constants in the upper and lower bound are independent of the choice of the stabilisation.

Remark 5.4 (Element-based estimator)**.** *The estimator η_ν is vertex-based. To obtain an element-based estimator, we can set*

$$\eta_K^2 := \sum_{\nu\in\mathcal{V}^K}\eta_\nu^2 \qquad\qquad \forall K\in\mathcal{T}_h.$$

The estimates in (5.10) also holds true for η_K if the summation happens on the elements in the upper bound, and the vertex patch ω^ν is replaced by the following element patch in the lower bound:

$$\omega^K = \bigcup_{\nu\in\mathcal{V}^K}\omega^\nu.$$

Remark 5.5 (Comparison with discontinuous Galerkin methods)**.** *As detailed in Lemmas 6.1 and 7.1 below, the estimator relies on the "broken" Prager–Synge estimation (5.4), instead of the usual Prager–Synge identity in (5.3).*

For discontinuous Galerkin schemes, the authors of [15] constructed two fields $\boldsymbol{\sigma}_h$ in $\boldsymbol{H}(\mathrm{div},\Omega)$ with $\mathrm{div}\,\boldsymbol{\sigma}_h = f$, and s_h in $H_0^1(\Omega)$ to insert in (5.3). Similar to ours, the construction in [15] for $\boldsymbol{\sigma}_h$ and s_h relies on vertex-patch finite element problems, with the crucial difference that these problems involve hat functions of the mesh as a partition of unity. Specifically, for each vertex ν in $\mathcal{V}_h$, the local finite element problems

$$s_h^\nu := \arg\min_{v_h\in\mathbb{P}_{p+1}(\mathcal{T}_h^\nu)\cap H_0^1(\omega^\nu)} \|\nabla_h(\varphi^\nu u_h) - \nabla v_h\|_{\omega^\nu} \tag{5.11a}$$

and

$$\boldsymbol{\sigma}_h^\nu := \arg \min_{\substack{\boldsymbol{\tau}_h \in \mathbb{RT}_p(\mathcal{T}_h^\nu) \cap \boldsymbol{H}_0(\mathrm{div},\omega^\nu) \\ \mathrm{div}\,\boldsymbol{\tau}_h = \varphi^\nu f - \nabla\varphi^\nu \cdot \mathfrak{G}_h(u_h)}} \|\varphi^\nu \mathfrak{G}_h(u_h) + \boldsymbol{\tau}_h\|_{\omega^\nu} \tag{5.11b}$$

are solved with suitable modifications at boundary vertices.

These local contributions are summed to obtain the globally conforming fields s_h and $\boldsymbol{\sigma}_h$. For all K in $\mathcal{T}_h$, the lower bound takes the form

$$\begin{aligned}&\|\mathfrak{G}_h(u_h) + \boldsymbol{\sigma}_h\|_K^2 + \|\mathfrak{G}_h(u_h) - \nabla s_h\|_K^2 \\ &\lesssim N_{\mathrm{vert}} C_{\mathrm{stab},K}^2 C_{\mathrm{cont},K}^2 \Big(\|\nabla u - \mathfrak{G}_h(u_h)\|_{\omega^K}^2 + \|\nabla_h u_h - \mathfrak{G}_h(u_h)\|_{\omega^K}^2 + \sum_{\substack{e \in \mathcal{E}_h \\ e \subset \overline{\omega^K}}} h_e^{-1} \|\Pi^0 [\![u_h]\!]\|_e^2 \Big),\end{aligned}$$

where

$$C_{\mathrm{stab},K} := \max_{\nu \in \mathcal{V}^K} C_{\mathrm{stab},\nu}, \qquad C_{\mathrm{cont},K} := \max_{\nu \in \mathcal{V}^K} C_{\mathrm{cont},\nu}.$$

A key asset of this construction is that the upper-bound in (5.1) *is constant-free, leading to guaranteed error bounds. However, the presence (and the computability) of the partition of unity in* (5.11) *is essential to obtain globally equilibrated fields s_h and $\boldsymbol{\sigma}_h$. As a result, this approach is fundamentally limited to simplicial and tensor-product meshes, where such closed-form partition of unity functions are available.*

In contrast, the broken approach we propose here enables to compute the estimator even if the partition of unity is not explicitly available, at the price of displacing the constants N_{vert} and $C_{\mathrm{cont},\nu}$ from the lower bound to the upper bound; see Theorem 5.3, and Lemmas 6.1 and 7.1 below. For this reason, the constants and terms in the upper and lower bounds here are a generalisation of those appearing in the Galerkin schemes in [15], and therefore they are as good as the state-of-the-art.

6 The flux-type term

We prove upper and (local) lower bounds for the second term on the right-hand side of (5.1).

Lemma 6.1 (Upper bound for the flux-type term)**.** *Assume that $\boldsymbol{\mathcal{G}}_h$ satisfies Assumption 5.1. Then, the following upper bound holds true:*

$$\sup_{\substack{v \in H_0^1(\Omega) \\ \|\nabla v\|_\Omega = 1}} \{(f,v)_\Omega - (\boldsymbol{\mathcal{G}}_h, \nabla v)_\Omega\}^2 \leq N_{\mathrm{vert}} \sum_{\nu \in \mathcal{V}_h} C_{\mathrm{cont},\nu}^2 \eta_{\nu,\mathrm{FL}}^2, \tag{6.1}$$

where the constant $C_{\mathrm{cont},\nu}$ appears in (4.5) *and N_{vert} denotes the maximal number of vertices of the elements of $\mathcal{T}_h$.*

Proof. We introduce the linear functional $\mathcal{R} \in (H_0^1(\Omega))'$ given by

$$\langle \mathcal{R}, v \rangle := (f,v)_\Omega - (\boldsymbol{\mathcal{G}}_h, \nabla v)_\Omega \qquad \forall v \in H_0^1(\Omega)$$

and its dual norm

$$\|\mathcal{R}\|_* := \sup_{\substack{v \in H_0^1(\Omega) \\ \|\nabla v\|_\Omega = 1}} \langle \mathcal{R}, v \rangle.$$

By Assumption 5.1, we have

$$\langle \mathcal{R}, \varphi^\nu \rangle = 0 \qquad \forall \nu \in \mathcal{V}_h^I. \tag{6.2}$$

Fix v in $H_0^1(\Omega)$, and $\zeta_\nu = 0$ if $\nu \in \mathcal{V}_h^B$ and $\zeta_\nu = (v,1)_{\omega^\nu}/|\omega^\nu|$ if $\nu \in \mathcal{V}_h^I$. Since the set $\{\varphi^\nu\}_{\nu \in \mathcal{V}_h}$ is a partition of unity, using (6.2), we have

$$\langle \mathcal{R}, v \rangle = \sum_{\nu \in \mathcal{V}_h} \langle \mathcal{R}, \varphi^\nu v \rangle = \sum_{\nu \in \mathcal{V}_h} \langle \mathcal{R}, \varphi^\nu (v - \zeta_\nu) \rangle.$$

Introducing the local dual norms

$$\|\mathcal{R}\|_{\nu,*} := \sup_{\substack{v\in H_0^1(\omega^\nu)\\ \|\nabla v\|_{\omega^\nu}=1}} \langle \mathcal{R}, v\rangle \tag{6.3}$$

and using (4.5) and the properties of ζ_ν, we arrive at

$$\langle \mathcal{R}, v\rangle \le \sum_{\nu\in\mathcal{V}_h} \|\mathcal{R}\|_{\nu,*}\, \|\nabla(\varphi^\nu(v-\zeta^\nu))\|_{\omega^\nu} \le \sum_{\nu\in\mathcal{V}_h} C_{\mathrm{cont},\nu}\, \|\mathcal{R}\|_{\nu,*}\, \|\nabla v\|_{\omega^\nu}.$$

Further using Cauchy-Schwartz' inequality leads to

$$|\langle \mathcal{R}, v\rangle|^2 \le \Big(\sum_{\nu\in\mathcal{V}_h} C_{\mathrm{cont},\nu}^2\, \|\mathcal{R}\|_{\nu,*}^2\Big)\Big(\sum_{\nu\in\mathcal{V}_h} \|\nabla v\|_{\omega^\nu}^2\Big) \le N_{\mathrm{vert}}\Big(\sum_{\nu\in\mathcal{V}_h} C_{\mathrm{cont},\nu}^2\, \|\mathcal{R}\|_{\nu,*}^2\Big)\|\nabla v\|_\Omega^2.$$

Since v is arbitrary, we conclude that

$$\|\mathcal{R}\|_*^2 \le N_{\mathrm{vert}} \sum_{\nu\in\mathcal{V}_h} C_{\mathrm{cont},\nu}^2\, \|\mathcal{R}\|_{\nu,*}^2. \tag{6.4}$$

So far, we proved an upper bound for the second term on the right-hand side of (5.1), i.e., we split it into the sum of local contributions. Next, we bound such local contributions on the right-hand side of (6.4) by means of a quantity that is computable via the degrees of freedom and can be obtained by means of local computations on vertex patches.

By Riesz' representation theorem, there exists r^ν in $H_0^1(\omega^\nu)$ such that

$$(\nabla r^\nu, \nabla v)_{\omega^\nu} = \langle \mathcal{R}, v\rangle \qquad \forall v\in H_0^1(\omega^\nu), \qquad\qquad \|\mathcal{R}\|_{\nu,*} = \|\nabla r^\nu\|_{\omega^\nu}. \tag{6.5}$$

The definition of $\mathcal{R}$ allows us to write

$$(\boldsymbol{\mathcal{G}}_h + \nabla r^\nu, \nabla v)_{\omega^\nu} = (f, v)_{\omega^\nu} \qquad\qquad \forall v\in H_0^1(\omega^\nu).$$

It follows that the function $\boldsymbol{\sigma}^\nu := -(\boldsymbol{\mathcal{G}}_h + \nabla r^\nu)$ belongs to $\boldsymbol{H}(\mathrm{div},\omega^\nu)$ with $\mathrm{div}\,\boldsymbol{\sigma}^\nu = f$. As a result, $\boldsymbol{\sigma}^\nu$ and r^ν are the solutions to the following mixed problem:

$$\begin{cases} \text{find } (\boldsymbol{\sigma}^\nu, r^\nu)\in \boldsymbol{H}(\mathrm{div},\omega^\nu)\times L^2(\omega^\nu) \text{ such that} & \\ (\boldsymbol{\sigma}^\nu,\boldsymbol{\tau}^\nu)_{\omega^\nu} + (\mathrm{div}\,\boldsymbol{\tau}^\nu, r^\nu) = (-\boldsymbol{\mathcal{G}}_h, \boldsymbol{\tau}^\nu)_{\omega^\nu} & \forall \boldsymbol{\tau}^\nu\in\boldsymbol{H}(\mathrm{div},\omega^\nu)\\ (\mathrm{div}\,\boldsymbol{\sigma}^\nu, q^\nu)_{\omega^\nu} = (f,q^\nu)_{\omega^\nu} & \forall q^\nu\in L^2(\omega^\nu). \end{cases} \tag{6.6}$$

Due to (6.3), (6.5), and (6.6), since $\boldsymbol{\sigma}^\nu$ solves the above mixed problem, we can write

$$\|\mathcal{R}\|_{\nu,*} = \|\boldsymbol{\sigma}^\nu + \boldsymbol{\mathcal{G}}_h\|_{\omega^\nu} = \min_{\boldsymbol{\tau}^\nu\in\boldsymbol{\Sigma}^\nu,\ \mathrm{div}\,\boldsymbol{\tau}^\nu = f} \|\boldsymbol{\tau}^\nu + \boldsymbol{\mathcal{G}}_h\|_{\omega^\nu}.$$

Let $\widetilde{\mathcal{T}}_h^\nu$ be the subtriangulation of ω^ν obtained by merging the subtriangulations $\widetilde{\mathcal{T}}_h^K$ of all elements K contained in ω^ν. For future convenience, we also define $\mathcal{T}_h^\nu$ as the set of elements K in $\mathcal{T}_h$ that are contained in ω^ν. Recalling the definition in (5.5), we arrive at

$$\|\mathcal{R}\|_{\nu,*} \le \min_{\boldsymbol{\tau}_h^\nu\in\mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu),\ \mathrm{div}\,\boldsymbol{\tau}_h^\nu = f} \|\boldsymbol{\tau}_h^\nu + \boldsymbol{\mathcal{G}}_h\|_{\omega^\nu} =: \eta_{\nu,\mathrm{FL}}.$$

The bound in (6.1) follows from (6.4). □

Next we show the (local) lower bound for the flux-type term.

Lemma 6.2 (Lower bound for the flux-type term)**.** *Let $C_{\mathrm{stab},\nu}$ be the constant in* (4.6b)*. For all ν in $\mathcal{V}_h$, we have*

$$\eta_{\nu,\mathrm{FL}} \le C_{\mathrm{stab},\nu}\|\nabla u - \boldsymbol{\mathcal{G}}_h\|_{\omega^\nu}.$$

Proof. The assertion is a consequence of the results stated in (4.6) and established in Appendix A. Indeed, since $\boldsymbol{\mathcal{G}}_h$ and $-\nabla u$ belong to $\mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu)$ and $H(\mathrm{div},\omega^\nu)$ with $\mathrm{div}(-\nabla u) = f$, respectively, we have

$$\eta_{\nu,\mathrm{FL}} := \min_{\substack{\boldsymbol{\sigma}_h^\nu\in\mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu)\cap H(\mathrm{div},\omega^\nu)\\ \mathrm{div}\,\boldsymbol{\sigma}_h^\nu = f}} \|\boldsymbol{\mathcal{G}}_h + \boldsymbol{\sigma}_h^\nu\|_{\omega^\nu} \le C_{\mathrm{stab},\nu} \min_{\substack{\boldsymbol{\sigma}^\nu\in H(\mathrm{div},\omega^\nu)\\ \mathrm{div}\,\boldsymbol{\sigma}^\nu = f}} \|\boldsymbol{\mathcal{G}}_h + \boldsymbol{\sigma}^\nu\|_{\omega^\nu} \le C_{\mathrm{stab},\nu}\|\boldsymbol{\mathcal{G}}_h - \nabla u\|_{\omega^\nu}.$$

□

7 The potential-type term

We prove upper and (local) lower bounds for the first term on the right-hand side of (5.1).

Lemma 7.1 (Upper bound for the potential-type term)**.** *For all $\boldsymbol{\mathcal{G}}_h \in \boldsymbol{L}^2(\Omega)$, the following upper bound holds true*

$$\begin{aligned}&\min_{s\in H^1_0(\Omega)} \|\boldsymbol{\mathcal{G}}_h - \nabla s\|_\Omega^2\\ &\le 4N_{\mathrm{vert}} \sum_{\nu\in\mathcal{V}_h} C_{\mathrm{cont},\nu}^2\Big(\eta_{\nu,\mathrm{PT}}^2 + \|\boldsymbol{\mathcal{G}}_h - \nabla_h\Pi_p^\nabla u_h\|_{\omega^\nu}^2 + \sum_{e\in\mathcal{E}_{h,\nu}} h_e^{-1}\|\Pi^0[\![\Pi_p^\nabla u_h]\!]\|_e^2\Big),\end{aligned}$$

where the constant $C_{\mathrm{cont},\nu}$ appears in (4.5) *and N_{vert} denotes the maximal number of vertices of the elements of $\mathcal{T}_h$.*

Proof. Recall that s_h^ν are the minimisers in the definition of $\eta_{\nu,\mathrm{PT}}$ in (5.5). For interior vertices, the mean value of the s_h^ν can be freely chosen; henceforth, we assume that $(s_h^\nu,1)_{\omega^\nu} = (\Pi_p^\nabla u_h,1)_{\omega^\nu}$ whenever ν belongs to $\mathcal{V}_h^I$.

We introduce

$$s_h := \sum_{\nu\in\mathcal{V}_h} \varphi^\nu s_h^\nu \in H^1_0(\Omega).$$

The triangle inequality entails

$$\min_{s\in H^1_0(\Omega)} \|\boldsymbol{\mathcal{G}}_h - \nabla s\|_\Omega^2 \le \|\boldsymbol{\mathcal{G}}_h - \nabla s_h\|_\Omega^2 \le 2\|\boldsymbol{\mathcal{G}}_h - \nabla_h\Pi_p^\nabla u_h\|_\Omega^2 + 2\|\nabla_h(\Pi_p^\nabla u_h - s_h)\|_\Omega^2.$$

As for the first term, we have

$$\|\boldsymbol{\mathcal{G}}_h - \nabla_h\Pi_p^\nabla u_h\|_\Omega^2 \le N_{\mathrm{vert}} \sum_{\nu\in\mathcal{V}_h} \|\boldsymbol{\mathcal{G}}_h - \nabla_h\Pi_p^\nabla u_h\|_{\omega^\nu}^2.$$

As for the second term, we plug in the partition of unity $\{\varphi^\nu\}_{\nu\in\mathcal{V}_h}$:

$$\|\nabla_h(\Pi_p^\nabla u_h - s_h)\|_\Omega^2 = \Big\|\sum_{\nu\in\mathcal{V}_h} \nabla_h(\varphi^\nu(\Pi_p^\nabla u_h - s_h^\nu))\Big\|_\Omega^2 \le N_{\mathrm{vert}} \sum_{\nu\in\mathcal{V}_h} \big\|\nabla_h(\varphi^\nu(\Pi_p^\nabla u_h - s_h^\nu))\big\|_{\omega^\nu}^2.$$

Next, we employ (4.5) and the fact that s_h^ν belongs to $H^1(\omega^\nu)$ with $(s_h^\nu,1)_{\omega^\nu} = (\Pi_p^\nabla u_h,1)_{\omega^\nu}$:

$$\|\nabla_h(\varphi^\nu(\Pi_p^\nabla u_h - s_h^\nu))\|_{\omega^\nu}^2 \le C_{\mathrm{cont},\nu}^2\Big(\|\nabla_h(\Pi_p^\nabla u_h - s_h^\nu))\|_{\omega^\nu}^2 + \sum_{e\in\mathcal{E}_{h,\nu}} h_e^{-1}\|\Pi^0[\![\Pi_p^\nabla u_h]\!]\|_e^2\Big).$$

The assertion follows from the triangle inequality:

$$\|\nabla_h(\Pi_p^\nabla u_h - s_h^\nu))\|_{\omega^\nu}^2 \le 2\|\boldsymbol{\mathcal{G}}_h - \nabla_h(\Pi_p^\nabla u_h)\|_{\omega^\nu}^2 + 2\|\boldsymbol{\mathcal{G}}_h - \nabla s_h^\nu\|_{\omega^\nu}^2.$$

□

Next we show the (local) lower bound for the potential-type term.

Lemma 7.2 (Lower bound for the potential-type term)**.** *Let $C_{\mathrm{stab},\nu}$ be the constant in* (4.6a)*. For all ν in $\mathcal{V}_h$, we have*

$$\eta_{\nu,\mathrm{PT}} \le C_{\mathrm{stab},\nu}\|\nabla u - \boldsymbol{\mathcal{G}}_h\|_{\omega^\nu}.$$

Proof. The assertion is a consequence of the results stated in (4.6) and established in Appendix A. Indeed, since $\boldsymbol{\mathcal{G}}_h$ and u belong to $\mathbb{P}_{p+1}(\widetilde{\mathcal{T}}_h^\nu)$ and $\widetilde{H}^1(\omega^\nu)$, respectively, we have

$$\begin{aligned}\eta_{\nu,\mathrm{PT}} &:= \min_{s_h^\nu\in\mathbb{P}_{p+2}(\widetilde{\mathcal{T}}_h^\nu)\cap\widetilde{H}^1(\omega^\nu)} \|\boldsymbol{\mathcal{G}}_h - \nabla s_h^\nu\|_{\omega^\nu}\\ &\le C_{\mathrm{stab},\nu} \min_{s^\nu\in\widetilde{H}^1(\omega^\nu)} \|\boldsymbol{\mathcal{G}}_h - \nabla s^\nu\|_{\omega^\nu} \le C_{\mathrm{stab},\nu}\|\boldsymbol{\mathcal{G}}_h - \nabla u\|_{\omega^\nu}.\end{aligned}$$

□

8 Numerical results

We present some numerical experiments. After introducing two test cases, in Section 8.1 we compare the accuracy of the generalised gradient $\mathfrak{G}_h(u_h)$ as opposed to the standard projected VEM solution $\nabla_h(\Pi_p^\nabla u_h)$. Next, we assess the performance of the adaptive scheme lead by the error estimator in (5.8). After discussing the general structure of the adaptive algorithm in Section 8.2, we show the performance of the h- and p-versions adaptive scheme in Sections 8.3 and 8.4, respectively.

Test cases. We consider two test cases. On the square domain $\Omega_1 := (0,1)^2$, we consider

$$u_1(x,y) := \sin(\pi\ x)\sin(\pi\ y).$$

On the L-shaped domain $\Omega_2 := (-1,1)^2 \setminus [0,1) \times (-1,0]$, given (r,θ) the usual polar coordinates at $(0,0)$, we consider

$$u_2(x,y) = u_2(r,\theta) := r^{\frac{2}{3}} \sin\left(\frac{2}{3}\theta\right).$$

8.1 An alternative error measure

Here, we are interested in comparing the performance of two different computable errors for method (3.3). Let u and u_h be the solutions to (2.1) and (3.3), and $\boldsymbol{\mathcal{G}}_h$ be the generalised gradient in (3.7). Given Π_p^∇ as in (3.1), the first quantity we are interested in is

$$\left\|\nabla u - \nabla_h(\Pi_p^\nabla u_h)\right\|_\Omega, \tag{8.1}$$

which is the standard error measure in virtual elements.

The second error is that defined in (5.3). Such an error can be expected to be different from that in (8.1), as $\boldsymbol{\mathcal{G}}_h$ contains also information on the chosen stabilisation. This is apparent from Tables 1 and 2, where we compare the two errors together with $\mathcal{E}(\Omega)$ in (5.9) for the exact solution u_1. We consider sequences of uniform hexahedral meshes, and take $p = 1$ and 4.

Table 1 Errors for the test case u_1 on a sequence of uniform hexahedral meshes with $p = 1$.

level	$\left\|\nabla u - \nabla_h(\Pi_p^\nabla u_h)\right\|_\Omega$	order	$\|\nabla u - \boldsymbol{\mathcal{G}}_h\|_\Omega$	order	$\mathcal{E}(\Omega)$	order
1	1.1892×10^{0}	–	1.1103×10^{0}	–	1.1529×10^{0}	–
2	6.4270×10^{-1}	0.89	5.9106×10^{-1}	0.91	6.2828×10^{-1}	0.88
3	3.3768×10^{-1}	0.93	3.0910×10^{-1}	0.94	3.1965×10^{-1}	0.97
4	1.7244×10^{-1}	0.97	1.5754×10^{-1}	0.97	1.6085×10^{-1}	0.99
5	8.7032×10^{-2}	0.99	7.9415×10^{-2}	0.99	8.0678×10^{-2}	1.00
6	4.3709×10^{-2}	0.99	3.9855×10^{-2}	0.99	4.0403×10^{-2}	1.00
7	2.1901×10^{-2}	1.00	1.9962×10^{-2}	1.00	2.0218×10^{-2}	1.00
8	1.0962×10^{-2}	1.00	9.9896×10^{-3}	1.00	1.0113×10^{-2}	1.00
9	5.4840×10^{-3}	1.00	4.9969×10^{-3}	1.00	5.0575×10^{-3}	1.00
10	2.7427×10^{-3}	1.00	2.4990×10^{-3}	1.00	2.5290×10^{-3}	1.00

Table 2 Errors for the test case u_1 on a sequence of uniform hexahedral meshes with $p = 4$.

level	$\left\|\nabla u - \nabla_h(\Pi_p^\nabla u_h)\right\|_\Omega$	order	$\|\nabla u - \boldsymbol{\mathcal{G}}_h\|_\Omega$	order	$\mathcal{E}(\Omega)$	order
1	7.2068×10^{-3}	–	7.1145×10^{-3}	–	7.2193×10^{-3}	–
2	1.2802×10^{-3}	2.49	1.3543×10^{-3}	2.39	1.4225×10^{-3}	2.34
3	7.8260×10^{-5}	4.03	9.1314×10^{-5}	3.89	1.0082×10^{-4}	3.82
4	4.3659×10^{-6}	4.16	5.3705×10^{-6}	4.09	6.2357×10^{-6}	4.02
5	2.5425×10^{-7}	4.10	3.2919×10^{-7}	4.03	3.9151×10^{-7}	3.99
6	1.5360×10^{-8}	4.05	2.0383×10^{-8}	4.01	2.4519×10^{-8}	4.00
7	9.4506×10^{-10}	4.02	1.2696×10^{-9}	4.00	1.5350×10^{-9}	4.00

The computation of $\boldsymbol{\mathcal{G}}_h$ can be seen as a cheap and local post-processing of the discrete solution leading to a comparable approximation of the exact gradient.

8.2 The adaptive algorithm

We consider the usual adaptive algorithm based on the four steps procedure

$$\textbf{SOLVE} \quad \Longrightarrow \quad \textbf{ESTIMATE} \quad \Longrightarrow \quad \textbf{MARK} \quad \Longrightarrow \quad \textbf{REFINE}.$$

Some details follow:

- while solving the virtual element method for a given mesh $\mathcal{T}_h$ and degree of accuracy p, there is no need to compute the generalised gradient $\boldsymbol{\mathcal{G}}_h$, as it is a tool for the computation of the error estimator, and is computed in a subsequent step as a postprocessing of u_h;
- while computing the local error estimator η_ν on each vertex patch ω^ν, it is necessary to compute the generalised gradient $\boldsymbol{\mathcal{G}}_h$, the vertex contributions $\eta_{\nu,\mathrm{FL}}$ and $\eta_{\nu,\mathrm{PT}}$, and the jumps of Π_p^∇;
- we use Dörfler's marking strategy with bulk parameter $\theta = 0.5$;
- one of the advantages of employing polytopic meshes in adaptive methods resides in easily handling hanging nodes and facets; since we are not interested in coarsening the mesh, we only employ initial uniform Cartesian and triangular meshes; h-refinements are performed by a standard splitting of each quadrilateral and triangular shape into four siblings with halved diameter, respectively.

8.3 The h-version

We are interested in assessing numerically the upper and lower bounds in Theorem 5.3. To this aim, given η and $\mathcal{E}(\Omega)$ as in (5.8) and (5.9), we introduce the effectivity index

$$\mathcal{I} := \frac{\eta}{\mathcal{E}(\Omega)}. \tag{8.2}$$

In Figure 2, we plot the effectivity index under h-uniform and adaptive mesh refinements with an initial uniform Cartesian mesh of 4 elements or the test case u_1 and 12 elements for the test case u_2, and check whether that remains constant for degrees of accuracy $p = 1$ and 2.

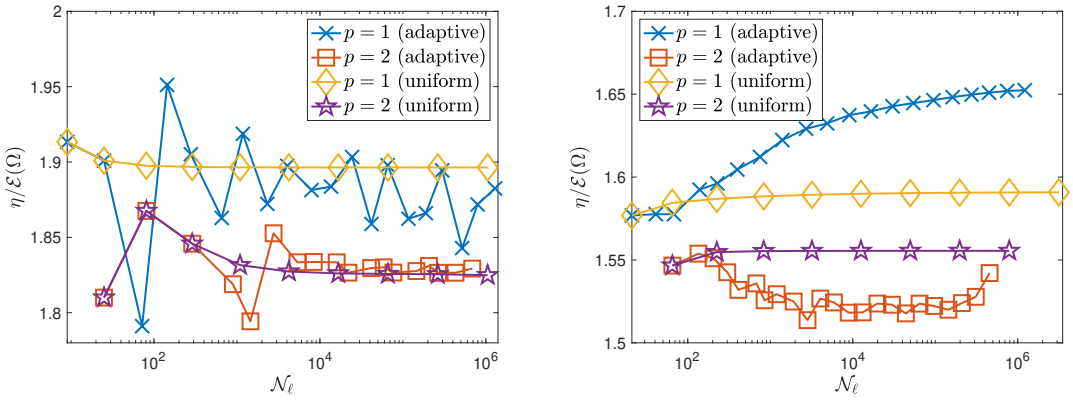

Figure 2: Effectivity index $\mathcal{I}$ in (8.2) versus the mnuber of degrees of freedom for the test cases u_1 (left panel) and u_2 (right panel) under uniform and adaptive h-refinements. We employ an initial uniform Cartesian mesh of 4 or 12 elements, respectively, and $p = 1$ and 2.

From Figure 2, it is apparent that the effectivity index remains in the bounded interval $(1.8, 3.3)$.

8.4 The p-version

We check the behaviour of the effectivity index $\mathcal{I}$ under uniform p-refinements, say, up to $p = 7$, on a uniform Cartesian and a uniform triangular meshes. We pick the exact solution u_2 and show the results in Figure 3.

Figure 3 illustrates the p-robustness of the proposed error estimator: the effectivity index remains in the bounded interval $(1.35, 1.6)$.

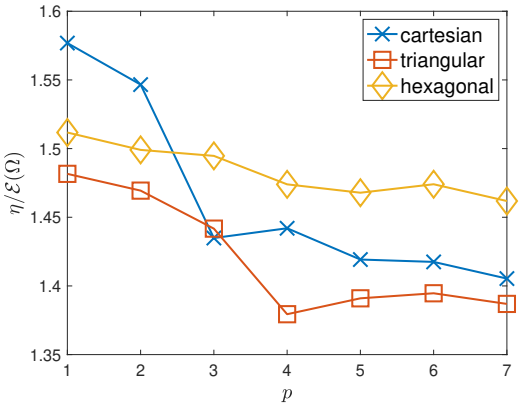

Figure 3: Effectivity index $\mathcal{I}$ in (8.2) for the test case u_2 under uniform p-refinements.

9 Conclusions and outlook

We proposed a generalised gradient formulation for the standard nodal virtual element method in two dimensions. On the one hand, the employed generalised gradient can be seen as a post-processing of the discrete solution, which provide us with an alternative computable approximation of the exact solution.

On the other hand, the generalised gradient formulation allowed us to improve the state-of-the-art of a posteriori error estimates in virtual elements and other polytopic methods in general. In fact, standard a posteriori error bounds on general meshes for arbitrary order degree of accuracy p typically involve constants depending on p and the choice of the stabilisations; the reason for this is to be sought both in the choice of the computable error measure and the error estimator. By changing these two quantities, for the first time in the literature on polytopic elements we ended up with robust a posteriori error estimates. The proposed error estimator contains terms and constants that extend those appearing while using other nonconforming methods, such as the discontinuous Galerkin method, on standard meshes.

Acknowledgments. LM has been partially funded by MUR (PRIN2022 research grant n. 202292JW3F). LM was also partially supported by the European Union (ERC Synergy, NEMESIS, project number 101115663). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. LM is member of the Gruppo Nazionale Calcolo Scientifico-Istituto Nazionale di Alta Matematica (GNCS-INdAM).

References

[1] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L.D. Marini, and A. Russo. Basic principles of virtual element methods. *Math. Models Methods Appl. Sci.*, 23(01):199–214, 2013.

[2] L. Beirão da Veiga, C. Canuto, R. H. Nochetto, G. Vacca, and M. Verani. Adaptive VEM: Stabilization-free a posteriori error analysis and contraction property. *SIAM J. Numer. Anal.*, 61(2):457–494, 2023.

[3] L. Beirão da Veiga and G. Manzini. Residual a posteriori error estimation for the virtual element method for elliptic problems. *ESAIM Math. Model. Numer. Anal.*, 49(2):577–599, 2015.

[4] L. Beirão da Veiga, G. Manzini, and L. Mascotto. A posteriori error estimation and adaptivity in *hp* virtual elements. *Numer. Math.*, 143:139–175, 2019.

[5] S. Berrone and A. Borio. A residual a posteriori error estimate for the virtual element method. *Math. Models Methods Appl. Sci.*, 2017.

[6] D. Braess, V. Pillwein, and J. Schöberl. Equilibrated residual error estimates are p-robust. *Comput. Methods Appl. Mech. Engrg.*, 198(13-14):1189–1197, 2009.

[7] S. C. Brenner. Poincaré–Friedrichs inequalities for piecewise H^1 functions. *SIAM J. Numer. Anal.*, 41(1):306–324, 2003.

[8] A. Cangiani, E. H. Georgoulis, T. Pryer, and O. J. Sutton. A posteriori error estimates for the virtual element method. *Numer. Math.*, 137(4):857–893, 2017.

[9] T. Chaumont-Frelet, A. Ern, and M. Vohralík. Stable broken H(curl) polynomial extensions and p-robust a posteriori error estimates by broken patchwise equilibration for the curl-curl problem. *Math. Comp.*, 91(333):37–74, 2022.

[10] T. Chaumont-Frelet and M. Vohralík. Constrained and unconstrained stable discrete minimizations for p-robust local reconstructions in vertex patches in the De Rham complex. `https://arxiv.org/abs/2208.05870`, 2022.

[11] F. Dassi, J. Gedicke, and L. Mascotto. Adaptive virtual elements with equilibrated fluxes. *Appl. Numer. Math*, 173:249–278, 2022.

[12] L. Demkowicz and M. Vohralík. p-robust equivalence of global continuous constrained and local discontinuous unconstrained approximation, a p-stable local commuting projector, and optimal elementwise hp approximation estimates in $\boldsymbol{H}(\mathrm{div})$. In preparation, 2024.

[13] Ph. Destuynder and B. Métivet. Explicit error bounds in a conforming finite element method. *Math. Comp.*, 68(228):1379–1396, 1999.

[14] D. A. Di Pietro, J. Droniou, and G. Manzini. Discontinuous skeletal gradient discretisation methods on polytopal meshes. *J. Comput. Phys.*, 355:397–425, 2018.

[15] A. Ern and M. Vohralík. Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous Galerkin, and mixed discretizations. *SIAM J. Numer. Anal.*, 53(2):1058–1081, 2015.

[16] A. Ern and M. Vohralík. Stable broken H^1 and $H(\mathrm{div})$ polynomial extensions for polynomial-degree-robust potential and flux reconstruction in three space dimensions. *Math. Comp.*, 89(322):551–594, 2020.

[17] L. Mascotto. The role of stabilization in the virtual element method: a survey. *Comput. Math. Appl.*, 151:244–251, 2023.

[18] J. M. Melenk and B. I. Wohlmuth. On residual-based a posteriori error estimation in hp-FEM. *Adv. Comput. Math.*, 15(1-4):311–331, 2001.

[19] I. Perugia, P. Pietra, and A. Russo. A plane wave virtual element method for the Helmholtz problem. *ESAIM Math. Model. Numer. Anal.*, 50(3):783–808, 2016.

[20] M. Vohralík. p-robust equivalence of global continuous and local discontinuous approximation, a p-stable local projector, and optimal elementwise hp approximation estimates in H^1. `https://inria.hal.science/hal-04436063`, 2024.

A Discrete stable minimisation

Given a vertex ν in $\mathcal{V}_h$, and data $\boldsymbol{G}_h$ and r_h in $\mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu)$ and $\mathbb{P}_p(\widetilde{\mathcal{T}}_h^\nu)$, we investigate the link between the discrete

$$m_h^{\mathrm{g}} := \min_{v_h \in \mathbb{P}_{p+2}(\widetilde{\mathcal{T}}_h^\nu) \cap \widetilde{H}^1(\omega^\nu)} \|\boldsymbol{G}_h - \nabla v_h\|_{\omega^\nu}, \qquad m_h^{\mathrm{d}} := \min_{\substack{\boldsymbol{\tau}_h \in \mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu) \cap \boldsymbol{H}(\mathrm{div},\omega^\nu) \\ \mathrm{div}\, \boldsymbol{\tau}_h = r_h}} \|\boldsymbol{\tau}_h + \boldsymbol{G}_h\|_{\omega^\nu} \tag{A.1}$$

and continuous minimisation problems

$$m^{\mathrm{g}} := \min_{v \in \widetilde{H}^1(\omega^\nu)} \|\boldsymbol{G}_h - \nabla v\|_{\omega^\nu}, \qquad m^{\mathrm{d}} := \min_{\substack{\boldsymbol{\tau} \in H(\mathrm{div},\omega^\nu) \\ \mathrm{div}\, \boldsymbol{\tau} = r_h}} \|\boldsymbol{\tau} + \boldsymbol{G}_h\|_{\omega^\nu}. \tag{A.2}$$

While it is clear that $m^{\mathrm{g}} \le m_h^{\mathrm{g}}$ and $m^{\mathrm{d}} \le m_h^{\mathrm{d}}$, the goal of this appendix is to show the converse inequality up to a constant. Specifically, in Section A.1 the inequality is shown for general meshes with constants possibly depending on p; in Section A.2, under more technical assumptions on the mesh, we prove that the constant does not depend on p.

A.1 Bounds for fully general meshes

Theorem A.1 (Non-p-robust stable minimisation)**.** *We have*

$$m_h^{\mathrm{g}} \le C(p,\gamma) m^{\mathrm{g}}, \qquad m_h^{\mathrm{d}} \le C(p,\gamma) m^{\mathrm{d}},$$

where the constants are independent of the data $\boldsymbol{G}_h$, r_h, and $\boldsymbol{G}_h$.

Proof. First, we establish the existence of generic constants C^{g} and C^{d} such that $m_h^{\mathrm{g}} \le C^{\mathrm{g}} m^{\mathrm{g}}$ and $m_h^{\mathrm{d}} \le C^{\mathrm{d}} m^{\mathrm{d}}$. We shall investigate the scaling properties in a second step. We only focus on C^{g}; the estimate involving C^{d} is established similarly.

Given s_h and s the discrete and continuous minimisers, the operators

$$A_h^{\mathrm{g}} : \boldsymbol{G}_h \in \mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu) \to \mathbb{P}_{p+1}(\widetilde{\mathcal{T}}_h^\nu) \ni \boldsymbol{G}_h - \nabla s_h$$

and

$$A^{\mathrm{g}} : \boldsymbol{G}_h \in \mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu) \to \boldsymbol{L}^2(\widetilde{\mathcal{T}}_h^\nu) \ni \boldsymbol{G}_h - \nabla s$$

are linear.

Hence, the following applications are seminorms on $\mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu)$:

$$\boldsymbol{G}_h \to \|A^{\mathrm{g}}(\boldsymbol{G}_h)\|_{\omega^\nu}, \qquad \boldsymbol{G}_h \to \|A_h^{\mathrm{g}}(\boldsymbol{G}_h)\|_{\omega^\nu}.$$

Since the space $\mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu)$ is finite dimensional, if we can show that $A_h^{\mathrm{g}}(\boldsymbol{G}_h) = 0$ whenever $A^{\mathrm{g}}(\boldsymbol{G}_h) = 0$, then these semi-norms have the same kernel. Thus, since all seminorms with the same kernel on finite dimensional spaces are equivalent, we may conclude that there exists a constant C^{g} such that

$$\|A^{\mathrm{g}}(\boldsymbol{G}_h)\|_{\omega^\nu} \le C^{\mathrm{g}} \|A_h^{\mathrm{g}}(\boldsymbol{G}_h)\|_{\omega^\nu}. \tag{A.3}$$

Estimate (A.3) would then imply that $m^{\mathrm{g}} \le C^{\mathrm{g}} m_h^{\mathrm{g}}$ for all $\boldsymbol{G}_h$ in $\mathbb{P}_{p+1}(\widetilde{\mathcal{T}}_h^\nu)$, which concludes the argument.

Therefore, we consider $\boldsymbol{G}_h$ in $\mathbb{RT}_p(\widetilde{\mathcal{T}}_h^\nu)$ such that $A^{\mathrm{g}}(\boldsymbol{G}_h) = 0$. We have that $\boldsymbol{G}_h = \nabla s$ in $\widetilde{H}^1(\omega^\nu)$ for the continuous minimiser s, so that $\boldsymbol{G}_h$ belongs to $\nabla(\mathbb{P}_{p+2}(\widetilde{\mathcal{T}}_h^\nu) \cap \widetilde{H}^1(\omega^\nu))$. It follows that $\boldsymbol{G}_h = \nabla s_h$ for the discrete minimiser too, leading to $A_h^{\mathrm{g}}(\boldsymbol{G}_h) = 0$.

The constants C^{g} and C^{d} obtained above depend on the space $\mathbb{P}_{p+1}(\widetilde{\mathcal{T}}_h^\nu)$, meaning that they depend on p and $\widetilde{\mathcal{T}}_h^\nu$. We can show that they only depend on $\widetilde{\mathcal{T}}_h^\nu$ through γ by using standard scaling arguments involving the Piola map and reference patches. □

A.2 Polynomial-degree-robust bounds

Here, under Assumption 2.1, we show that the discrete stable minimisation property holds true with a constant independent of p.

We start by considering the case of interior vertices.

Lemma A.2 (p-robust stable minimisation for interior vertices)**.** *Let Assumption 2.1 hold true. Assume that ν is an interior vertex and* (2.2a) *is satisfied. Then, we have*

$$m^{\mathrm{g}} \le C(\gamma) m_h^{\mathrm{g}}, \qquad m^{\mathrm{d}} \le C(\gamma) m_h^{\mathrm{d}}.$$

Proof. For patches consisting of simplices sharing a single vertex, this result is now standard; see, e.g., [6, 10, 16]. However, here, as the polygons are further broken down into additional triangles, the patches we consider do not enter this framework: some triangles do not have ν as a vertex. Nevertheless, a careful read of [10, 16] reveals that the proof should still go through for the above patches, as long as we can provide a suitable ordering (or enumeration) of the triangles in the patch.

In fact, the properties required from the enumeration have recently been explicitly stated in [20, Definition B.1]. Under the assumption that a suitable enumeration is available, the bound $m^{\mathrm{g}} \le C(\gamma) m_h^{\mathrm{g}}$ is explicitly established in [20, Theorem D.1]. The corresponding bound for the divergence-constrained problem can similarly be obtained following [10,16] with the aforementioned enumeration, and the details should appear shortly in [12].

Therefore, we have to produce an enumeration of the patch triangles meeting the requirements of [20, Definition B.1], i.e., list the elements as $\widetilde{\mathcal{T}}_h^\nu = \{T_1, \dots, T_N\}$ in such a way that:

a) for $1 < j \le N$, the triangle T_j shares at least on edge with an already enumerated triangle, i.e., there exists T_ℓ, $1 \le \ell < j$ such that $\partial T_j \cap \partial T_\ell \neq \emptyset$;

b) for $1 \le j \le N$, the triangle T_j shares at most two faces with already enumerated triangles;

c) for $1 \le j \le N$, if T_j shares two faces with, say T_ℓ and $T_{\ell'}$, $1 \le \ell < \ell' \le j$, then all the elements sharing the vertex common to T_j, T_ℓ and T_ℓ' are already enumerated.

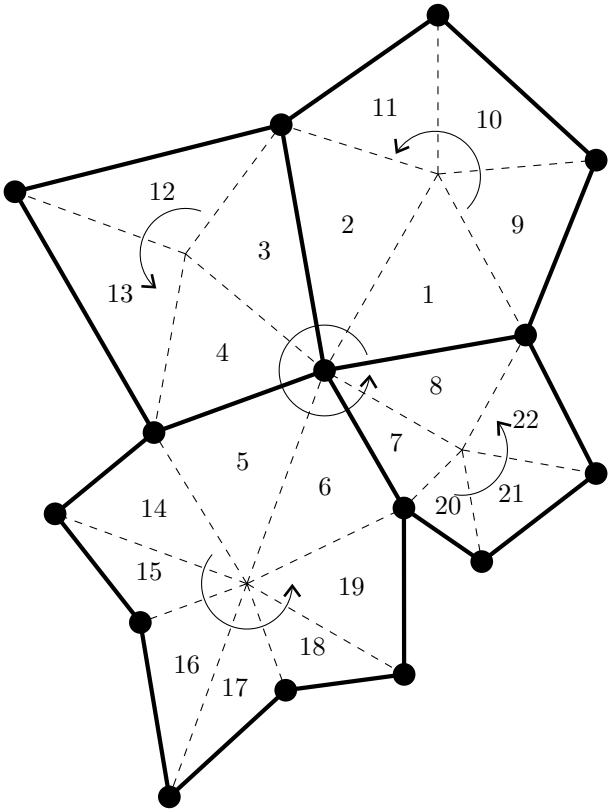

Figure 4: Enumeration in an interior vertex patch

For the standard case where all the triangles share the same vertex, i.e., as in [6], such an enumeration can be trivially constructed by looping around the vertex. Here, the enumeration can be obtained as follows, see Figure 4 for an illustration: first, we enumerate the triangles sharing the vertex ν counter clockwise; we then go through the centroid $\mathbf{x}_K$ of the polygonal elements K in $\mathcal{T}_h^\nu$ counter clockwise. Around each $\mathbf{x}_K$, we there is a closed patch of triangles sharing $\mathbf{x}_K$ and covering K, and two of them, say T_ℓ and $T_{\ell+1}$, have already been enumerated when looping over ν. Besides, due to (2.2a), the remaining triangles around $\mathbf{x}_K$ do not share faces with any triangle from another polygon K' in $\mathcal{T}_h^\nu$. We can then enumerate those triangles by completing the loop around $\mathbf{x}_K$ counter clockwise, i.e., going from T_ℓ to $T_{\ell+1}$ by crossing faces. □

Remark A.3. *If the triangles T_9 and T_{22} shared a face in Figure 4, then the enumeration would not be valid anymore. Indeed, all the neighbours of T_{22} would have appeared sooner in the enumeration, hence violating point c) in the proof above. This explains why we invoke* (2.2a) *to carry out the proof.*

Nevertheless, (i) the presence of elements K sharing multiple faces does not exclude the existence of a suitable enumeration satisfying [20, Definition B.1], even though admittedly, we were not able to provide a generic construction; moreover, (ii) the lack of a suitable enumeration does not imply that p-robustness does not hold in general. In other words, the question of p-robustness on general patches is still open. We deem that there is no severe obstruction to its proof and that it could be obtained at the price of more technicalities in the construction of a suitable enumeration.

Eventually, we deal with the exterior vertices case.

Lemma A.4 (p-robust stable minimisation for boundary vertices)**.** *Let Assumption 2.1 hold true. Assume that ν is a boundary vertex and* (2.2b) *is satisfied. Then, we have*

$$m^{\mathrm{g}} \leq C(\gamma) m_h^{\mathrm{g}}, \qquad m^{\mathrm{d}} \leq C(\gamma) m_h^{\mathrm{d}}.$$

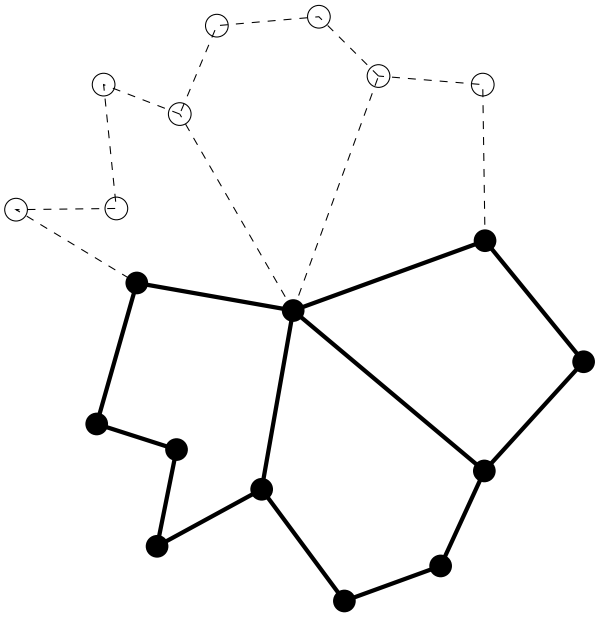

Figure 5: Symmetrisation of boundary patch

Proof. We only sketch the main ideas of the proof, since it closely follows that of [16, Section 7] and [10, Section 7]. The key idea is to symmetrise the patch around the boundary composed of the exterior faces sharing the vertex ν to obtain an interior patch. This process is illustrated in Figure 5. Display (2.2b) ensures that the boundary actually only consists of two faces, and the symmetrised patch satisfies (2.2a). The symmetrised patch is an interior patch, for which Lemma A.2 now applies; the arguments developed in [10, 16] can be invoked to show that, if the discrete stable minimisation property holds true in the symmetrised patch, then it must hold true in the original patch too. □

Collecting the estimates above, we end up with the following result.

Theorem A.5 (p-robust stable minimisation)**.** *Let Assumption 2.1 hold true. Assume that Assumption 2.1 is satisfied. Then, we have*

$$m^{\mathrm{g}} \leq C(\gamma) m_h^{\mathrm{g}}, \qquad m^{\mathrm{d}} \leq C(\gamma) m_h^{\mathrm{d}}$$

where the constant $C(\gamma)$ is independent of p.